The Measurement of Verbal and Numerical Abilities

BF
21
A7
no.107-

Schneck

Class 151.22
Book No. 40955

Northeastern University
Library

THE MEASUREMENT OF VERBAL AND NUMERICAL ABILITIES

BY
MATTHEW M. R. SCHNECK, Ph.D.

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, Editor
No. 107

NEW YORK
June, 1929

Psych
BF
21
A7
no. 107

ACKNOWLEDGMENTS

For suggestion and definition of the problem, for the general plan of the investigation, and for almost daily consultation and guidance with respect to procedure and statistical treatment, I am deeply indebted to Professor Henry E. Garrett, of the Department of Psychology, Columbia University. Without Professor Garrett's direction, this study could not have been made. To the Department of Psychology, Columbia University, I extend thanks for relieving me of the task of calculating many correlations. The Department was kind enough to finance this work, which was done by the Columbia University Statistical Bureau. I am also indebted to Professor Howard D. Marsh of the Department of Philosophy, College of the City of New York, for permission to use his classes in Experimental Psychology, and for valuable cooperation in other directions. I am grateful for the kindness of Professor Egbert Turner, who turned over to me his classes in Educational Psychology at the College of the City of New York, for preliminary experimentation. My thanks are also extended to Mrs. Georgette Decsi Schneck for valuable assistance in the scoring of the tests and in calculations.

THE AUTHOR

1929

6 Feb 1952

TABLE OF CONTENTS

Chapter		*Page*
I.	The Problem	5
II.	The Literature	7
III.	The Procedure	15
	1. The Subjects	15
	2. The Tests	15
	3. The Testing	18
	4. The Scoring	20
IV.	The Results	21
	1. Means, Standard Deviations and Intercorrelations	21
	2. Statistical Analysis and Discussion	23
	3. Relation of Verbal and Numerical Abilities to Other Abilities	40
	4. Relation of Verbal and Numerical Abilities to Scholastic Records	42
V.	Summary	48
VI.	Bibliography	49

40955

The Measurement of Verbal and Numerical Abilities

I. The Problem

This paper is a specific contribution to an increasing body of literature dealing with the organization of mental traits. The center of interest for the present study lies in the interrelationships existing between two manifestations of intellectual function, the term 'intellect' being used in much the same sense as that in which it is used by Thorndike (22). These functions may be termed, for want of more precise names, *verbal ability* and *numerical ability*. The writer hesitates to assign names of any kind to these activities, for the use of substantives to denote activities is always hazardous, and almost certain to give rise to needless controversy. The designations will not be insisted upon, nor will they be defended. It is desirable, nevertheless, to explain what is meant by the terms employed. For the purpose of this paper, then, verbal ability will refer simply to performance in those tests which involve, predominantly, a knowledge of words and ability to use them. The essence of the performance demanded by those tests which will be called verbal is the possession by the subject of the accomplishments just mentioned. Similarly, numerical ability will refer to performance in those tests which involve, predominantly, the ability to manipulate number and number concepts. Verbal and numerical tests are designated as such in the body of this report.

It is not pretended that exhaustive tests of either kind of ability have been made, nor that complete measurements of 'verbality as such' or of 'numerical ability as such' have been achieved. An attempt was made, indeed, to secure a fairly representative sampling of these two abilities. Whether or not the attempt has succeeded is a matter which the writer prefers to submit to the judgment of the reader. This, however, is insisted upon. Whatever the abilities involved *in the tests* may be, it is with these abilities that the writer is concerned. Since definitions in psychology are sometimes less than definitive, it is felt to be better merely to designate rather than to define.

If, then, it be permitted to refer to the ability or abilities called upon in the verbal tests here employed as *verbal ability,* and to the ability or abilities involved in the numerical tests as *numerical ability,* the following questions will be dealt with:

1. Are verbal ability and numerical ability discrete and independent capacities, or do they possess elements in common? 2. Are there group factors of verbal ability and of numerical ability, either or both?* 3. Are measurements of these abilities merely measurements of ability in general? 4. What is the best single test or battery of tests for the measurement of these abilities? 5. What is the relationship between verbal and numerical ability, on the one hand, and, on the other hand, performance in college courses presumably involving these abilities?

* In the system of Spearman, a group of abilities may be divided into a single general factor which is common to all of the abilities, plus specific factors, each of which is present in one, but not more than one, ability. This is true provided a certain statistical criterion, to be described later, is satisfied. If specific factors overlap, so that they are present in more than one, but not in all of the abilities, the criterion cannot be satisfied, and a group factor, caused by the overlapping, is present.

II. The Literature

A search for literature bearing directly upon the foregoing problems yields scant return. There have been, indeed, a host of measurements of verbal ability and of numerical ability, but few of these studies have gone beyond a mere calculation of coefficients of correlation, from the magnitude of which extremely dubious conclusions have been drawn. The popular notion has been that verbal intelligence is something quite different from mathematical intelligence, and that an individual may be proficient in the one, and, at the same time, deficient in the other. This notion gains some support from a study of the correlations reported. These are nearly always positive, but rarely high. Thorndike (23) has said that "The ability measured by verbal tests is not the same as the ability measured by non-verbal tests." This seems to be the prevailing notion, though Spearman holds a contrary view. According to Spearman (17) all mental performances partake of one general common function called 'g', and both verbal and non-verbal tests really measure this function.

C. Burt (2, p. 46ff) administered a series of tests to 120 children "of the same sex, differing but little in age, zeal, attendance and social status," and found the following special abilities: arithmetical, manual, linguistic and composition. The linguistic group, according to Burt, is not entirely unrelated to the composition group, and the two may perhaps be taken as forming a single literary group. Burt's method of locating these special abilities is of doubtful accuracy, depending as it did upon an inspection of his correlation tables. He observed that the values found in this table did, to a certain extent, conform to a theoretical hierarchical arrangement, which indicated the presence of a common general factor. In some cases, however, the correlations are far too high; and these cases he takes to be indicative of the presence of special factors additional to the universal factor common to all of the subjects. Now, the method of inspection of a table to determine the presence or absence of hierarchy is decried by Spearman himself (17, p. 138). (For the significance of the hierarchy, see Spearman, (17)). A better approximation is obtained by the use of the intercolumnar correlation coefficient (17), but even this has been rejected by Spearman in favor of the cri-

terion of the tetrad difference. (17, p. 79). It may be added that the present writer's inspection of Burt's correlation table yields conclusions which are by no means in agreement with those of Burt. It appears to the writer that no particular arrangement of any kind is present, and that neither a universal factor nor special factors can be inferred with safety.

A more recent study is that of Davey (6). As a pupil of Spearman, in whose laboratory the problem of group factors is a very live issue, Miss Davey undertook the comparison of ability in verbal tests with ability in tests which were of similar form, but non-verbal. She constructed a battery of pictorial tests similar in form to a series of oral tests. Her subjects were 243 children in London County Council schools, ranging in age from 8 to 14 years, and divided into ten groups. Six of these groups were in different classes of different schools. A seventh group consisted of 82 boys, all in the same school, divided into 4 sub-groups according to age. The total number of subjects of each sex is not given, nor is the number of subjects in each of the groups. It is obvious, however, that her groups must have been very small. There was no mixture of the sexes in any of the groups, but there may have been diversity in age, excepting in the 4 sub-groups of boys, mentioned above. Intercorrelations among the tests given were calculated for each of the ten groups, and these intercorrelations averaged. The average intercorrelations seem to have a satisfactory degree of reliability, but it is regrettable that the individual coefficients were not published. An average of ten correlations may be itself reliable when the single correlations, or some of them, are unreliable. Working with such small groups, it is very probable that some of the correlations were unreliable. The propriety of working with such average correlations may be questioned.

Eight oral and six pictorial tests were given to each of the groups, and the correlations calculated by means of the Spearman Rank-order formula. Averaging the intercorrelations, it appears that the oral and pictorial tests correlate to about the same extent as do the pictorial tests among themselves, but that the correlation for the oral tests *inter se* is somewhat higher. The P. E. of the difference in the latter case (calculated by the present writer) is not completely significant, if 4 P. E. be accepted as the significant criterion. Miss Davey, of course, does not accept these data as definitive evidence of

anything. She calculated 420 tetrad differences, and found that the median tetrad difference is .021, with a probable error of .0194, which she interprets as indicative of the presence of a group factor. This view is rather surprising, coming from Spearman's laboratory, for in the latter's volume, *The Abilities of Man,* a median tetrad difference which is not much larger than its probable error is uniformly regarded as insignificant.

Analyzing the table of tetrad differences, Miss Davey concludes that there is no group factor in the pictorial tests; that there is a group factor in the oral tests; but that this group factor is confined to the first four of these oral tests. From this result, she comes to the further conclusion that there is no group factor of verbality as such, for if there were, it would have shown itself in all eight tests. Just what is meant by 'verbality as such' is not made entirely clear, though some light is thrown upon the matter by a further analysis, as the result of which Miss Davey concludes that the group factor was introduced by the content of the tests, rather than by their form. Even if we grant the validity of the method by which this conclusion was arrived at (and it is by no means granted), there still appears to be a distinction without much difference. Form as distinguished from content is altogether a matter of interpretation. The distinction made here seems to be another attempt to reify an abstract concept. To the present writer, this situation seems to justify the notion of Wilson (24) that 'g' is relative to the set-up, unless one can get a situation in which $r_{ag} = 1$*. It is probably true that suitable changes in the content of Miss Davey's first four oral tests might have eliminated the group factor which she found. This would prove merely that it is possible to devise tests which are so very different from one another that no group factor is present. It would prove nothing about 'verbality as such'. Granted that the form of the tests is not responsible for the introduction of the group factor; granted that the responsibility lies with the content, or with the similarity of the relations educed; granted further, that a change in content would eliminate the group factor; we must necessarily conclude that, since the form does not change, content is responsible both for the group factor when it is present, and for its absence when it is absent. All of which seems to reduce form, so far as group factors are

* r_{ag} = the correlation between the general factor, 'g' and any given ability, 'a'.

concerned, to insignificance. Verbality as such seems to be dependent upon the set-up, and to have no meaning whatever apart from the set-up. Miss Davey's final conclusions are to the effect that a verbal test measures the same 'g' as does a test similar in form but non-verbal. The conclusion may express a truth, but it certainly does not follow from Miss Davey's results. At least, her own verbal tests measure some group factor which is absent from her pictorial tests.

A. I. Gates (10) used verbal and non-verbal material for the purpose of predicting success in school subjects. Twenty third-grade subjects were given seven non-verbal and five verbal tests. The average intercorrelations follow:

Verbal	.62
Non-verbal	.40
Verbal and non-verbal	.24

From these results Gates concludes that the verbal tests clearly yield results different from the non-verbal. While it is unsafe to draw conclusions from data obtained from so small a group, the results point in the direction of the prevailing opinion.

T. L. Kelley (15), working with students in the seventh and third grades and in the kindergarten, has located group factors as follows: verbal, number, memory, spatial and speed. He has, in addition, devised a method for measuring these group factors. In the same volume (p. 192ff) Kelley analyses the data reported in Rose G. Anderson's monograph (1). Rearranging and combining Anderson's results, and applying the earlier steps of his own procedure, Kelley locates a verbal, a numerical and a memory bond.

In general, a verbal factor seems to be indicated in all of the studies cited. By reason of the criticisms advanced above, the evidence appears to be less than conclusive. Even Kelley admits that his groups were not to his liking so far as homogeneity is concerned. This matter of homogeneity of groups is a vexing one. Groups composed of individuals differing materially from one another in age, nurture, nationality, social and economic conditions, and the like, are dangerous to work with, for the correlations obtained may be spurious. Thus, if a group of boys and girls, in equal number, and of the same age, be measured for strength of grip and for general intelligence, and the correlations between these capacities be cal-

culated, it is obvious that the correlation will be spurious. The boys will be high, mediocre and low in intelligence, as will, also, the girls. But the boys, as a body, will rank higher in strength of grip than will the girls. In such a case, the correlation between strength of grip and general intelligence is very likely to be zero, or even negative. The same condition will hold if the correlation were calculated between general intelligence and height for a group composed of Japanese and Swedes. The correlation will certainly be lower than it would be if the group were composed entirely either of Japanese or of Swedes. Again, Kelley (15, p. 25) has shown how a correlation may be raised by the introduction of diversity in maturity. So that diversity within a group may raise a correlation, or it may lower it.

The distorting influence of diversity has been stressed by Kelley (15) and by Spearman (17, p. 155), the former insisting that heterogeneity tends to introduce correlation, the latter contending that heterogeneity will reduce it. Either result may follow, as has been shown above. In view of Spearman's insistence upon homogeneity, it is curious that two of the principal studies upon which he relies for the maintenance of his thesis were, in this respect, glaring offenders. Thus, Spearman makes much of the study of Bonser (17, pp. 139, 147, xii, xx). But Bonser's group of 757 subjects was composed of 385 boys and 372 girls, ranging in age from 8 to 16, in the 4th, 5th and 6th grades of five public schools. This is heterogeneity with a vengeance. Again, Spearman stresses the study of Simpson (17, p. 145). For these data, Spearman calculated 3003 tetrad differences, and hails the result as one of the most striking agreements between theory and observation ever found in psychology, adding that this agreement would not easily be matched in any other science. But if ever a group was heterogeneous, it is precisely the group upon which these momentous conclusions are based. Simpson worked with a total of 37 subjects, which is obviously a very small group. Worse than this, the subjects were 17 advanced students and professors, and 20 adults of the sort which may almost be termed, appropriately, human derelicts. It is difficult to imagine a greater degree of diversity than just this group. Even if it be agreed, with Spearman, that heterogeneity reduces correlation, it is surely a strange procedure to make use of a widely diverse group merely because correlation happens

to be present. It may very well be that the correlation is present just because the group is heterogeneous. There will be occasion to advert to this matter later.

In addition to the verbal factor which has been indicated, the studies of Burt, Kelley and Anderson point to the existence of a numerical factor. The evidence, on the whole, is meagre. A few studies do publish correlations between arithmetic and verbal tests. Thus, Bonser (4), working with children in the fourth, fifth and sixth grades, finds a correlation of +.41 between arithmetic and completion tests, and a correlation between arithmetic and opposites amounting to +.42. Whether these data do or do not indicate special factors is not determined. Indirect evidence of more or less independence of traits is furnished by studies of special disabilities. Bronner, for example, finds marked disability for language in cases where all other processes are normal. She reports similar clinical findings for numerical ability. (5, Chaps. 5 and 6). Special disabilities of this sort are also reported by Symonds (19), Hollingworth and Winford (13), C. Burt (3) and A. I. Gates (11).

The problem of group factors has assumed major importance in recent years, largely as a result of the assiduous labors of Spearman and his pupils. One study from that laboratory has already been reviewed above. In general, Spearman (17, Chap. 13) is not greatly impressed by the evidence for group factors. After analyzing this evidence, he states that "cases of specific correlations or group factors have been astonishingly rare." He does find evidence for group factors "in respect of what may be called the logical, the mechanical, the psychological, and the arithmetical abilities." He acknowledges also a special ability for appreciation of music. A slight degree of confusion appears in Spearman's account, for he employs the term "group factor" in more than one sense. Thus, he recognizes perseveration, oscillation and persistence of motives as universal factors, but in several places he also refers to them as group factors. The latter, in Spearman's system, are defined as "those which occur in more than one but less than all of any given set of abilities." It is in this sense that the writer will speak of group factors in this paper.

Since Spearman's statistical technique is, to a certain extent, employed in this report, it seems desirable to devote some space to it. It is not deemed necessary to trace the development

of the technique, since this has already been done by Dodd (7, 8). Briefly, then, Spearman holds that when certain statistical conditions are satisfied, it is possible to divide cognitive abilities into two parts, independent of each other. The one part is a *general* factor, called 'g'. It varies from individual to individual, but remains the same for a given individual in respect of all of the correlated abilities. The other part is *specific*, varying from individual to individual, and, for any given individual, from ability to ability. The statistical condition which must be satisfied is as follows:

If a group of individuals is tested in a number of abilities, the intercorrelations between all possible pairs of the variables will, if arranged in the usual correlation table, furnish the necessary data for the application of the method. If we let 'a' stand for one variable, 'b' for a second, and 'p' and 'q' for a third and fourth, it can be shown that the abilities tested are divisible into one general factor and four specific factors when

$$r_{ap} \times r_{bq} - r_{aq} \times r_{bp} = 0$$

throughout the table. The entire expression is called the tetrad equation, and the right side of the equation is termed the tetrad difference. When four variables are used, three tetrad equations are calculated. With a greater number of variables, the number of tetrad equations to be calculated is given by ${}_3{}_n C_4$. Fortuitous influences are always present, so that the tetrad equation will rarely equal zero exactly. It must be shown, then, that the probable error of the tetrad difference is large enough to make the difference insignificant. When this can be shown for all of the tetrad differences in the table, then the abilities in question can be divided as above. When tetrad differences are significantly greater than 0, we have evidence of the presence of a group factor or factors which Spearman interprets as the overlapping of specific factors. The proof of these propositions is given in (17, Appendix, iii).

The system sketched above has been attacked very vigorously, notably in a series of papers by G. H. Thomson (21), who presents some trenchant criticisms of Spearman's interpretations, especially of 'g'. The present writer is, in this paper, very little concerned with the nature of 'g'. For his purposes, the divergent views of both Spearman and Thomson are quite immaterial. The tetrad criterion does indicate the

presence of something which is shared by all of the variables, whatever that may be. Further than this one need not venture, though the writer will have occasion to offer some remarks bearing upon this central factor.

III. The Procedure

1. *The Subjects.* Reference was made above to the desirability of working with a group of subjects in which diversity is reduced in so far as this is possible. The writer was fortunate in having, in his own classes, such a group. All of the subjects used were students in a class in General Psychology, at the College of the City of New York. The tests were given to the entire class, and the unsuitable subjects eliminated after the testing. Those retained numbered 210. All of them were men of Hebrew extraction, whose parents were of foreign birth. The group was further limited to those whose parents were born in Russian, Polish or Balkan countries. A few of the subjects were sons of professional men, but these were eliminated, so that the final group included only those whose fathers were engaged in manual trades or in small retail establishments of their own management. Seven upper sophomores and lower seniors are included, the remaining 203 subjects being in the junior year. All of the subjects were educated in the New York City public schools, including the elementary school, the high school and the college, the latter being a municipal institution, tuition free. Nearly all were self-supporting, in whole or in part. In age, the subjects ranged from 18 to 21. This introduces diversity, but diversity which is less damaging than would be a similar age range in the earlier years. Furthermore, age was partialled out before any of the significant statistical steps were taken, so that in this respect the group was homogeneous. To summarize: The group is fairly large, and quite similar with respect to education, ancestry, economic and social status and age. All were of the male sex. On the whole, heterogeneity was reduced to a minimum.

2. *The Tests*: Nine tests in all were given to all of the subjects. Five of these were intended to measure verbal ability, and four were intended to measure numerical ability. The tests were of the following types:

A. *Verbal Tests.**

1. *Vocabulary Test:* This contained 130 items, of the selective type, in which the person tested was required to choose,

* The writer will be glad to supply copies of the tests to interested persons.

from a group of four words, that one whose meaning was most nearly the same as that of the test word. The instructions were as follows:

"Look at the first word in Line 1. Find the other word in the line which means the same or most nearly the same as the first word. Underscore this synonymous word.

helotry diabolism gambling slavery assistance

The word which is most nearly the same as helotry is slavery, which is therefore underscored. Do the same in each line below. Underscore one word in each line and only one."

About 30 words were taken from the CAVD test and about ten from the Detroit Advanced Intelligence Test, as well as a few from the Ohio State University Intelligence Test. In most cases, the writer found it necessary to make changes in the words from which choice was to be made, in order to fit the test for a group of upperclass college men. The greater bulk of the test was composed by the writer. The net result was a test of unusual difficulty, but not too difficult for this group.

2. *Opposites Test:* This test was of the same form and length as the Vocabulary Test, and of equal difficulty. The instructions were similar, except that the testees were required to underscore the word which was most nearly the opposite of the key word. A few of the words in this test were taken from the Ohio State University Intelligence Test, the rest being prepared by the writer.

3. *Analogies Test:* This consisted of 40 items, ranging from easy to very difficult items. The instructions were as follows:

"In each of the lines below you will observe the relationship between the first and second words. From among the words contained in the parentheses, choose the word which bears the same relation to the third word of the line as the second word bears to the first. Thus:

Sky : blue :: grass: (table green warm big)

Here the word green, contained within the parentheses, bears the same relationship to 'grass' as 'blue' does to 'sky'. The word 'grass' is therefore underlined. Do the same in each of the lines below. Underline one word in each line, and only one word. Remember: The word to be underlined is always in the parentheses." Eighteen of the analogies were taken from Army Alpha and from the Miller Mental Ability Test. The remainder, comprising the more difficult items, were prepared by the writer.

4. *Sentence Completion Test:* This was composed of 50 items, all taken from the CAVD Test, Levels M, N, O, P, Q. The instructions given were those used in the published test

5. *Disarranged Sentences:* This test contained 40 items, all prepared by the writer. Following are the instructions given:

"Read the sentences as they would be if the words were arranged in correct order. Do not rewrite the sentences. When the sentence asks a question, answer it on the blank line. When no question is asked, do what the sentence tells you to do."

B. *Numerical Tests:*

1. *Arithmetic Reasoning Test:* This test was composed of 40 items, of various types. The items were taken from the CAVD Test; from the Otis Self-Administering Test of Mental Ability, Higher Examination, Form A; from Morgan's Mental Test, Form A; from the Brown University Psychological Examination, Series II, Exercise F; and from the Rogers Test of Mathematical Ability. A few were prepared by the writer.

2. *Number Series Completion:* This test contained 40 items. Four were taken from Army Alpha, the remainder being composed by the writer. The instructions were:

"Look at each row of numbers below. On the two dotted lines write the two numbers that should come next, as in the samples."

Samples:

2	4	6	8	10	12	14	16
9	8	7	6	5	4	3	2

3. *Equation Relations:* This test contained 26 items, of two forms. Sixteen of the items were of the type given in the CAVD Test, some of them being drawn directly from that test, the rest being added by the writer. The instructions for these problems were:

"In the problems below, write the numbers and signs in the proper order, so that they make a true equation. The letters in the problem stand for arithmetical signs, as shown in the key. Whenever a letter appears, refer to the key and convert the letter into the proper mathematical sign."

Key: A stands for the plus sign.
B stands for the minus sign.
C stands for the multiplication sign.
D stands for the division sign.
E stands for the equals sign.
F stands for parentheses, i.e., ( ).

Samples: Problem	*Solution*
3 3 6 E A	$3 + 3 = 6$
4 7 8 20 E A C	$7 \times 4 = 20 + 8$

Ten items were of the form given below. The instructions follow:

"In the problems below, only the equality sign is given. It is indicated by the letter E, as above. You are to supply the signs needed, and to write the numbers and signs in the correct order, so that they make a true equation. The numbers which you find at the left of the E will remain on that side in the finished equation. Those at the right of the E will remain on the right. You are simply to rearrange the numbers on each side of the E, and to supply the missing signs, to make a true equation."

Thus: 2 7 E 3 18 3 Solution: $7 + 2 = 18 - (3 \times 3)$

These items were all prepared by the writer. They proved to be somewhat less difficult than the items of the other form.

4. *Mental Multiplication:* This test contained 58 items, composed by the writer. The instructions follow:

"In each row you will find a multiplicand and a multiplier, as indicated by the column headings. Multiply *mentally* the multiplicand by its accompanying multiplier, and write the result in the column marked 'Product.' Thus, in the first row, the multiplicand is 14; the multiplier is 11. Multiply the 14 by the 11 and write the result on the dotted line to the right of the 11. Do the same for each row. *Remember:* do no figuring with pencil and paper. All calculations must be done mentally, without any aids."

It is acknowledged that the tests could have been improved. This is particularly true of the test of Disarranged Sentences, which proved to have been somewhat too easy. Scores for this test were piled up at the upper end of the curve. The Analogies Test should have been longer, for its reliability turned out to be somewhat lower than was desired. In general, the tests were fairly satisfactory, so far as reliability is concerned. (See Results.) Only the Analogies and the Disarranged Sentences failed to meet the standard set, which was a reliability of .90 or better. The reliability of the Analogies was .879, and the Disarranged Sentences only .755.

3. *The Testing:*

All of the tests were administered by the group method. The subjects were given mimeographed sheets which contained the

test material. The Analogies, Disarranged Sentences and Opposite tests were given in two groups of 120 each. The remainder of the tests were given in the laboratory, to groups of 18. An attempt was made to have the items of the tests arranged in order of difficulty. This order was determined by preliminary testing of students in classes in Educational Psychology. None of these students were in the experimental group itself. It was found that the order of difficulty determined in this way was maintained fairly well in the experimental group. Difficulty was determined by the number of failures for each item.

With the exception of the test in Mental Multiplication, all tests were power tests. Multiplication was a speed test. The time allowed was as follows:

Vocabulary	35 minutes
Opposites	35 minutes
Analogies	30 minutes
Disarranged Sentences	25 minutes
Sentence Completion	50 minutes
Arithmetic Reasoning	45 minutes
Number Series	30 minutes
Equation Relations	50 minutes
Multiplication	10 minutes

The total testing time was, therefore, *5 hours and ten minutes.* The testing was distributed over three weeks, it being thought undesirable to harass the students unnecessarily. It is believed that the subjects entered into the program in a spirit of cooperation. The tests were sufficiently difficult to constitute a genuine intellectual challenge. In the case of the student-body of the College of the City of New York, this is all that is necessary to insure interest and enthusiasm. Few of the students had ever been tested before, so that it was not found necessary to deal with a condition of negative adaptation to a testing program. The work was done at a time when the intelligence-testing movement was under discussion in class, and the program, therefore, fitted naturally into the general scheme of things. The members of the class were not told the purpose of the experiment, and most of them did not realize that an experiment was going forward. They were told, however, that the results of the testing would have no effect upon their term grades. In general, then, the unnatural conditions of experimentation were minimized. The time allotted for the tests proved to be sufficient, so that 90% of the subjects were

able to finish the work, except, of course, in the case of the Multiplication. Some of the men did finish this test in the ten minutes allowed, but most did not. None, however, failed to complete at least 30 items in Multiplication.

IV. *The Scoring:* All tests excepting the Sentence Completion were scored in terms of the number of items done correctly. The sentences were scored on a scale ranging from 0 to 3, the latter being allowed for completions admitted as correct by the key used for the CAVD Test at Teachers' College, Columbia University. The score of 0 was given in cases of complete failure, or when the completion did not make good sense. The scores of 1 and 2 were allowed for cases in which the completion did not agree with the key, but in which there was not failure. When only one word differed from the key, the score of 2 was given. When more than one differed, the item was allowed one point. In both cases, of course, the sentence as completed had to make reasonably good sense. Since this scheme was followed for all subjects, it is felt that the results of scoring by this method were valid.

The Disarranged Sentences were so constructed that only one correct response was possible. It was observed that, in nearly all cases, the response was either correct or there was no response at all. One item of the Arithmetic Test admitted of two correct answers, and both were accepted. This was the first item in the test. In the Number Series Completion Test, it was required that both dotted lines be filled in correctly. If either was incorrect, the whole item was counted as incorrect. The Equation Relations and the Multiplication presented no difficulties. In the Multiplication, the answer was either correct or it was not. In the Equations Relations, only those equations which were correct throughout were accepted. No partial credit was given. The correct responses in the Vocabulary, Opposites and the Analogies Tests were determined by reference to manuals of synonyms and antonyms, and the words from which selection was to be made were so chosen as to permit only one correct response.

In general, then, the scoring was done with a high degree of objectivity. Subjective judgment was, in some degree, inevitable. Actually, individual judgment is never absent altogether, even in physical measurements. It is less prominent in some situations than in others, but it cannot be wholly avoided.

IV. The Results

1. *Means, Standard Deviations and Intercorrelations*

Means, standard deviations and intercorrelation coefficients were computed by the Columbia University Statistical Bureau, all coefficients being calculated by machine and verified carefully. The reliability coefficients were calculated by the writer. The odd-numbered items were correlated against the even-numbered items, and the result corrected by the application of the Spearmen Brown Prophecy Formula (9, p. 269, Formula No. 59), to give the reliability of a test of double length. The means, standard deviations and reliability coefficients are given in Table I. The reliability coefficients will be found printed in italics, and lying along the diagonal from the first to the tenth variable. The tenth variable is age, whose correlation with each of the tests was calculated, age being stated in months. The inclusion of age as a variable was deemed desirable for the purpose of determining its effect upon the results, by the use of the partial correlation technique.

It will be observed that in all cases except one, that of Mental Multiplication, the correlation with age is negative. This is quite consistent with the usual findings. Within a given grade, the younger students are frequently of superior ability, and this seems to be true in the junior year in college. The negative correlations are low, but they seem to be fairly uniform in all cases excepting Multiplication. Here the correlation is positive and very slightly above zero. Actually, it may be taken as an absence of correlation. The reason for this exceptional value for Multiplication and Age is not entirely clear. It may be conjectured that, in such a relatively mechanical operation as multiplication, growth in ability approaches its limit at or before the age of 18, so that there is little differentiation to be found as the result of maturity or of precocity. It may be that the superiority of the younger members of a grade or college year is evidenced in respect of the relatively involved mental operations, those requiring greater reasoning ability and relation-perceiving ability, than in respect of the relatively easy or mechanical operations. If this be true, the absence of negative correlation would be explained, for multiplication makes less demand upon "intellectual" pro-

TABLE I

INTERCORRELATIONS AND RELIABILITY COEFFICIENTS OF NINE TESTS AND OF AGE, TOGETHER WITH THE MEANS AND S.D.'S OF THE TESTS

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1. Vocabulary	*.906*	.8818	.6878	.5362	.3020	.2711	.1827	.0476	.0575	—.2377
2. Oppositese		*.925*	.7170	.5374	.2646	.2889	.1778	.0539	.0458	—.2081
3. Analogies			*.879*	.4496	.2857	.2478	.1660	.0540	.0356	—.1727
4. Sentence Completion				*.921*	.2573	.3619	.3966	.1806	.1140	—.2883
5. Disarranged Sentences					*.755*	.1161	.0403	.0701	—.0255	—.2166
6. Arithmetic Reasoning						*.920*	.4517	.3754	.3728	—.2236
7. Number Series Completion							*.907*	.2512	.3083	—.2399
8. Equation Relations								*.923*	.2702	—.2497
9. Multiplication									*.954*	.0087
10. Age										*1.0000*
Mean	73.85	80.20	27.37	64.69	32.87	21.21	31.29	13.53	38.62	235.0
S.D.	15.74	15.15	4.18	20.07	3.95	6.25	6.73	3.70	11.16	11.3

Note: The Mean and S.D. of Age are given in months.

cesses than do the operations involved in the other tests. The writer does not insist upon this explanation, but he does regard it as probable, ruling out, of course, cases of phenomenal arithmetical ability as well as those cases in which some special scheme of mental multiplication is employed. Such schemes do exist, and the writer has no means of knowing how many of his subjects availed themselves of these short-cuts. A short-cut of this type would certainly tend to reduce even further the 'intellectual' elements in multiplication.

The reliability coefficients are, on the whole, satisfactory. All of the numerical tests have a reliability of .90 or more. Two of the verbal tests fall below this standard, but even in these cases the reliability is fairly high. The reason for the deficiencies of these tests has been suggested above.

It is possible to obtain a first impression of the relation between the verbal and the numerical tests by an inspection of the raw coefficients. The average intercorrelations are as follows:

Verbal Tests	.4920
Numerical Tests	.3383
Verbal & Numerical Tests	.1441

These results may be compared with those of Gates, cited in the foregoing. Gates, also, found that the average intercorrelations for the verbal and non-verbal tests was lower than those for either the verbal or non-verbal *inter se*. The evidence at this point is not definitive, but already it begins to point to a substantial difference between the verbal and numerical tests. The analysis must, of course, be carried much farther before final conclusions may be drawn.

2. *Statistical Analysis and Discussion*

A. It would have been desirable to confine the experimentation to subjects who were all of the same age. Unfortunately, this was not possible, for the size of the group would have been reduced quite drastically. Since age was included as a tenth variable, and its correlation with each of the other variables computed, it becomes possible to eliminate the influence of age by statistical treatment. If, then, age be partialled out, we have remaining the intercorrelations of the tests with the age factor held constant. This was done, with the results shown in Table II. Since the correlations with Age were, in all cases but one, negative, the effect of holding Age constant is to re-

duce the intercorrelations. Only where Multiplication is involved are the correlations raised, since the correlation of this variable with Age was positive.

B. The next step involves the correction of the coefficients in Table II for chance variations or errors of observation. Each of the coefficients was therefore corrected for attenuation, using the formula applicable when only one test for each variable is available. (9, Formula No. 48.) These corrected coefficients will be employed hereafter in the calculation of tetrad differences. According to Spearman (17, Appendix vi), the tetrad difference criterion is free from the influence of attenuation. "If the criterion is passed when the correlations are corrected for attenuation, then it must also be passed when they are not so." The writer does not dispute the validity of Spearman's mathematical demonstration of this proposition, *when the tetrad difference equals zero.* But tetrad differences rarely equal zero exactly. Spearman regards such a difference as insignificant when the tetrad difference is less than four or five times its probable error. It is quite possible, however, that a difference which, by Spearman's rather exacting standards, is insignificant, may become quite significant when corrected for attenuation. Conversely, it is conceivable that a tetrad difference which seems to be significant will become insignificant when the coefficients are corrected.

A coefficient of correlation is corrected for attenuation, by the method used in this study, by dividing the coefficient by the square root of the product of the reliability coefficients involved. Obviously, when the right side of the equation is zero, this division will not alter it. When, as is nearly always true, the tetrad equation does not equal zero the right side must be divided in the same manner as the left side, and by the same value. Now, the magnitude of the correction for attenuation depends upon the value of the reliability coefficients. The greater the value of the reliability coefficients, the smaller will be the corrections. If, then, the corrections to the right side of the minus sign in the tetrad equation be greater than those at the left of the minus sign, the tetrad difference becomes smaller than it would be in the absence of correction. If, on the other hand, the corrections to the left of the minus sign be greater, the tetrad difference will be larger than it would be without correction for attenuation. In this study, two tetrad differences which were statistically significant before correc-

TABLE II

INTERCORRELATIONS OF NINE VARIABLES WITH AGE PARTIALLED OUT

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. Vocabulary	....	.8760	.6760	.5028	.2651	.2302	.1333	—.0125	.0614
2. Opposites		....	.7070	.5097	.2308	.2543	.1347	.0082	.0487
3. Analogies			....	.4239	.2589	.2179	.1303	.0114	.0377
4. Sentence Completion				....	.2097	.3187	.3522	.1171	.1164
5. Disarranged Sentences					....	.0722	—.0011	.0018	—.0014
6. Arithmetic Reasoning						....	.4207	.3386	.3845
7. Number Series Completion							....	.2035	.3197
8. Equation Relations								....	.2813
9. Multiplication									....

tion became insignificant after correction. It seems necessary, therefore, to correct coefficients for attenuation before the application of the tetrad criterion. Conclusions, otherwise, may be seriously erroneous. The corrected values of the coefficients of Table II are given in Table III.

These 'true' coefficients seem to indicate even more clearly than did the uncorrected data, the presence of a real difference between the verbal and the numerical tests. The average correlations follow:

Verbal	.5238
Numerical	.3507
Verbal and Numerical	.1223

Both for the verbal tests *inter se* and for the numerical tests *inter se*, the average intercorrelation is higher than before correction for attenuation. On the other hand, the average intercorrelation between the verbal and the numerical tests is lower. These data tell us nothing concerning the presence or absence of factors peculiar to either the verbal or the numerical abilities. Whether these factors exist must be determined by other means. We proceed, therefore, to the application of Spearman's technique, in so far as this technique is useful for present purposes.

C. *The Tetrad Difference Criterion*

When a correlation table contains nine variables, the number of tetrad differences to be calculated is 378. If each of these differences is zero, within the limits of its probable error, then each of the nine variables may be thought of as containing one general factor, shared by all of the variables, plus a specific factor or factors peculiar to each of the variables. The present problem, may, however, be approached somewhat differently. If our correlation table be divided into two parts, the one containing only the intercorrelations of the verbal tests, and the other only the intercorrelations of the numerical tests, we may calculate the tetrad differences for each of these two tables. It will then be possible to compute the correlation between the central 'verbal' factor and the central 'numerical' factor, provided the tetrad difference criterion be satisfied in both tables. This procedure eliminates much of the mechanical labor entailed by the calculation of 378 tetrad differences.

It will be observed (Table III) that the correlation between

TABLE III

INTERCORRELATIONS FOR NINE VARIABLES WITH AGE PARTIALLED OUT AND AFTER CORRECTION FOR ATTENUATION

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. Vocabulary	....	.9569	.7575	.5504	.3206	.2521	.1471	—.0137	.0660
2. Opposites		....	.7841	.5522	.2762	.2757	.1471	.0089	.0518
3. Analogies			....	.4711	.3178	.2423	.1459	.0127	.0412
4. Sentence Completion				....	.2515	.3462	.3854	.1270	.1242
5. Disarranged Sentences					....	.0866	—.0013	.0022	—.0016
6. Arithmetic Reasoning						....	.4605	.3674	.4104
7. Number Series Completion							....	.2224	.3437
8. Equation Relations								....	.2998
9. Multiplication									....

Vocabulary and Opposites is .9569. This correlation is very much higher than are any of the others in the table, so high, indeed, that the two tests may be taken to involve identical abilities. This conclusion is supported by reference to the tests themselves. They are of the same form and of nearly equal difficulty. The only difference in the abilities called for is that in the one a relation of similarity is to be observed, while in the other the relation is that of opposites. It was decided, therefore, to eliminate one of these tests. An alternative to elimination would have been Spearman's procedure of pooling the results of the two tests. Since the judgment concerning the essential similarity of the two abilities is, in part, subjective, it was felt that the sounder procedure was the one decided upon. Either test might have been eliminated, but it was decided to eliminate Opposites, since the tetrad differences obtained by its inclusion were somewhat higher than when the Vocabulary test was used. It should be remarked, however, that in both cases the tetrad differences satisfy the criterion, as may be seen from the data given in Table IV. In this table, tetrad differences are designated in the manner adopted by Kelley (15, p. 11).

TABLE IV

TETRAD DIFFERENCES OBTAINED FROM TABLE III

	A. Verbal Tests	
	Opposites Eliminated	*Vocabulary Eliminated*
	T.D. P.E.	T.D. P.E.
t_{1234}	.0156 ± .0308	.0217 ± .0323
t_{1243}	.0395 ± .0281	.0671 ± .0287
t_{1342}	.0239 ± .0174	.0454 ± .0162
	B. Numerical Tests	
	T.D. P.E.	
t_{1234}	.0118 ± .0238	
t_{1243}	.0468 ± .0226	
t_{1342}	.0350 ± .0202	

In no case is the tetrad difference more than three times its probable error. The probable errors were calculated by the use of Kelley's formula* for the probable error of a tetrad

$$*P.E.t_{1234} = .6745 \times \frac{1}{\sqrt{N}}[r_{12}^2 + r_{13}^2 + r_{24}^2 + r_{34}^2 + 2r_{12}\,r_{14}\,r_{23}\,r_{34} + 2r_{13}\,r_{14}\,r_{23}\,r_{24} - 2r_{12}\,r_{13}\,r_{23} - 2r_{12}\,r_{14}\,r_{24} - 2r_{13}\,r_{14}\,r_{34} - 2r_{23}\,r_{24}\,r_{34} + t_{1234}^2\,(r_{12}^2 + r_{13}^2 + r_{14}^2 + r_{23}^2 + r_{24}^2 + r_{34}^2 - 4)]^{\frac{1}{2}}$$

difference (15, p. 49, Formula 28). This formula differs from the one offered by Spearman and Holzinger (17, Appendix, xi). Kelley points out the fact that Spearman and Holzinger's formula applies when the true tetrad difference is zero, whereas his own formula does not depend upon this. When the true difference is zero, then Kelley's formula becomes identical with that of Spearman and Holzinger. Holzinger (14) remarks that Kelley has retained a term which, in the actual calculation is negligible. This is true, but it was nevertheless deemed safer to use the longer formula, especially since the added labor is not irksome when only a few tetrad differences are calculated. Since each of our tables contains only four variables, it was not necessary to compute more than three tetrad differences for each. If, then, we accept the standard which requires that the tetrad difference be at least four times its probable error to be considered significant, it appears that the criterion is satisfied. It is probable that this standard is too exacting, but it seems best, at this point, to meet Professor Spearman on his own ground. It is concluded, then, that our verbal tests may be thought of as consisting of a central factor which is present in all of the verbal tests, plus a factor or factors specific to each of the tests, and not shared by any of the others. A similar conclusion is drawn for the numerical tests. The central factor for the verbal tests will hereafter be referred to as V, and that for the numerical tests as N. In neither case is any hypothesis advanced concerning the nature of these central or common factors.

D. *Correlations between V and N*

Since the major purpose of this study is the determination of the relation of verbal abilities to numerical abilities, it seems desirable to compute the correlation between V and N. Obviously, the specific factors involved have no place in this scheme since, by their very specificity, their correlations are zero, within the limits of their probable errors. If, then, there is any correlation between verbal and numerical abilities, it must consist in the correlation of their common 'general' factor or factors. The calculation of this correlation is arrived at by a circuitous route, to be detailed in the following pages.

1. The first step is the determination of the correlation between each of the verbal tests and V, as well as the correlation of each of the numerical tests with N. For this purpose, Spear-

man's formula (17, Appendix, Formula 21) was employed, yielding the results shown in Table V.

TABLE V

CORRELATIONS OF NINE TESTS WITH THEIR CENTRAL FACTORS, V OR N

	Opposites Eliminated	*Vocabulary Eliminated*
A. Verbal Tests		
Vocabulary and V	.8965	
Opposites and V		.8781
Analogies and V	.8160	.8435
Sentence Completion and V	.6081	.6128
Disarranged Sentences and V	.3841	.3622
B. Numerical Tests		
Arithmetic Reasoning and N	.7666	
Number Series Completion and N	.5593	
Equation Relations and N	.4613	
Multiplication and N	.4867	

The correlations with V differ somewhat, depending upon whether Vocabulary or Opposites be retained. These differences for the last three verbal tests should be well within the limits of the probable errors of these differences. These probable errors, then, were calculated, with the following results:

TABLE VI

PROBABLE ERRORS OF THE DIFFERENCE OF THE CORRELATIONS OF THE VERBAL TESTS WITH V AS DEPENDENT UPON THE ELIMINATION OF EITHER THE VOCABULARY OR OPPOSITES TESTS

	Obtained Difference	*P.E. of Difference*
Analogies and V	.0275	.0205
Sentence Completion and V	.0047	.0413
Disarranged Sentences and V	.0219	.0566

In no case, then, is the difference significant, as may be seen by a comparison of each of the obtained differences with its probable error.

The magnitude of the correlation coefficient of Table V is an index of the extent to which the several tests measure the common factors. The Vocabulary Test appears to be the best instrument for the measurement of that which is common to all of the verbal tests. Since the correlations have been corrected for attenuation, the reliability of the Vocabulary Test is 1.00. Its correlation with V is virtually .90, which is not far below its correlation with itself. Using uncorrected coefficients, the correlation of this test with V is .84, while its

obtained reliability is .91. A test whose correlation with a criterion is not far below its correlation with itself is a very satisfactory instrument for the measurement of that criterion. It should be recalled, in this connection, that Professor Terman (20) uses a vocabulary test in his revision of the Binet Scale as a preliminary determinant of the lower limit of the testee's mental age. He finds a knowledge of the meaning of words the best single criterion for this purpose. This is indirect evidence that 'general intelligence' as measured by the Stanford Binet Scale and other tests is largely verbal ability, or the ability to manipulate language.

The Analogies Test is somewhat less satisfactory as a measure of V, though its correlation with V is still above .80. This might have been expected, since this test also involves knowledge of words in much the same sense as does the Vocabulary Test. That it contains other elements, however, is indicated by the fact that the corrected correlation of Analogies with Vocabulary is only .76; and this fact leads to the conclusion that there is something more in V than mere ability to select synonyms.

The test in Sentence Completion is much less satisfactory than are Vocabulary and Analogies, while Disarranged Sentences seem to be distinctly unsatisfactory. It is the least reliable of the tests, and its distribution is noticeably skewed. Whether the shortcomings of these two tests lie in their form or in their content is not determined. Certainly their forms are quite different from those of the Vocabulary and Analogies tests. Their content must naturally be different, otherwise we would have a series of identical tests. Whatever these tests measure, they assuredly do not measure the same thing in equal degree.

None of the numerical tests is quite as satisfactory as is the Vocabulary Test or the Analogies. The highest correlation is that between Arithmetic Reasoning and N. Its magnitude is .7666. The other numerical tests are distinctly inferior to Arithmetic, so far as efficiency of measurement of N is concerned. No single member of our battery of numerical tests can be used as an approximately sufficient index of N. Whatever the nature of N, it appears that a considerable portion of it is left untouched by our numerical tests. This is true even when the entire battery is correlated with N, as will be shown presently.

2. It is possible to calculate the correlation of V with any combination of the verbal tests, to determine which battery will give the best measure of V. This can be done, also, for N and the numerical tests. These correlations are arrived at by finding the necessary multiple correlation coefficients. Thus, the expression $R_{1.2345}$ denotes the correlation between one variable and four other variables taken as a team. If, then, the tests be designated by numbers as shown in the following table, the multiple correlations may be found.

TABLE VII

CORRELATION OF *V* AND *N* WITH EVERY POSSIBLE COMBINATION OF TESTS

See Key Below

	Verbal	*Numerical*
$R_{1.2345}$	.9306	.8294
$R_{1.234}$	.9281	.8209
$R_{1.235}$	.9240	.8205
$R_{1.245}$	.9119	.8060
$R_{1.345}$	.8590	.6953
$R_{1.23}$	.9207	.8011
$R_{1.24}$	.9070	.7906
$R_{1.25}$	.9023	.7895
$R_{1.34}$	.8545	.6574
$R_{1.35}$	.8266	.6412
$R_{1.45}$	.6533	.5833

Key: $_1$ = V & N
$_2$ = Vocabulary & Arithmetic
$_3$ = Analogies & No. Ser. Completion
$_4$ = Sent. Compl. & Equations
$_5$ = Dis. Sent. & Multiplication

It is obvious that the best measure of both V and N is obtained with the use of the full batteries. When the last two tests of the verbal battery are used alone, the result is far from satisfactory. This is consistent with the data of Table V, where it was found that the correlation of the central factor with these tests, when taken singly, is relatively low. In the verbal set, the multiple correlation is reduced, whenever the Vocabulary Test is excluded, to a greater degree than when only the Analogies Test is omitted. This, again, is in accord with Table V. The same tendencies are observable in the numerical tests. Omission of the Arithmetic Test from any combination causes a sharp drop in the size of the multiple correlation. This test, taken alone, had the highest correlation with N. The numerical battery, taken as a unit, does not measure its central factor as efficiently as the verbal battery

measures its central factor. The numerical battery is better than the best single numerical test when taken alone, but it leaves something to be desired.

Whatever is common to the numerical tests is, in each of the single tests, apparently diluted by extraneous factors which are washed out when the group is taken as a unit. The nature of these extraneous factors is, at present, open to speculation. The Arithmetic Test includes the use of verbal elements, which may operate to reduce the correlation of this test with the 'numerical factor.' The other numerical tests also demand some verbal operations, but to a lesser degree than does the Arithmetic. It would seem, then, that the other tests ought to have a higher correspondence with N than does Arithmetic. The opposite situation, is, however, true. It may be, then, that the arithmetic test, which has a greater variety of operations than do the others, does by this very fact of variety include more of 'numerical ability.' Taken alone, then, the efficiency of this test is reduced by verbal factors. The addition of the other tests may result in the contribution of other 'numerical elements' which have the effect of offsetting the extraneous elements. They may, in addition, bring to bear other 'numerical elements' which are not present in the Arithmetic Test, over and above those which are required to nullify the effects of the verbal elements. It is admitted that this hypothetical explanation is a sheer guess, one which the writer does not care to defend very vigorously.

3. *Correlations of Tests with Specific Factors*

The mathematical derivation of the tetrad difference criterion is predicated upon the theory of partial correlation (17, Appendix iii). In this derivation, it follows from the definition of r_{ap} that $r_{ap.g} = 0$.* This means that when the tetrad difference criterion is satisfied, there can be no correlation between the specific factors contained in the variables. Such correlation as exists between the variables must be due entirely to 'g', the general factor which is shared by all of the variables. Since it is possible to find the correlation of each of the variables with 'g,' it must also be possible to find the correlation of each variable with its own specific factor or factors. This is

* R_{ap} designates the correlation between two variables, 'a' and 'p.' $R_{ap.g}$ is the correlation between 'a' and 'p' with the factor which is common to both eliminated or held constant. $R_{ap.g.}$, then, is the correlation between the factors which are specific to 'a' and 'p,' respectively.

accomplished by determining to what extent the variable is not correlated with 'g'. This follows necessarily from the fact that, when the criterion is satisfied, each variable is divisible into two parts, the general and the specific. Now, the extent to which any variable is not correlated with 'g' is found, after its correlation with 'g' has been computed, by substituting the latter value in the formula for the 'coefficient of alienation' (16, p. 173). For this purpose we have available the data of Table V, from which the following specific correlations are obtained:

TABLE VIII

CORRELATIONS OF NINE VARIABLES WITH THEIR SPECIFIC FACTORS

Vocabulary	.4430
Analogies	.5781
Sentence Completion	.7939
Disarranged Sentences	.9233
Arithmetic Reasoning	.6421
No. Series Completion	.8290
Equation Relations	.8872
Multiplication	.8736
Verbal Battery	.3667
Numerical Battery	.5587

Most of these correlations are sufficiently high to be quite impressive. Despite the fact that the verbal battery has a correlation of .93 with V, there is nevertheless a high degree of specificity in two of these tests, and a substantial degree of specificity in the others. This is even more noticeable in the numerical battery. In three of these tests, the specific ingredient overshadows the general factor markedly, so far, at least, as the correlations are concerned.

4. *V and N Scores*

Knowing the correlation of each of the tests with its central factor, we may now take the next step toward the determination of the correlation of V with N. It is necessary to find a V score and an N score for each of the subjects of the experiment. Since the obtained scores are weighted unequally, they cannot be combined into a composite score directly. It is first required that the weight of each separate score be reduced to 1.00. Such weighted scores are known as 'reduced' scores. Woodworth's method of weighting was used in computing reduced scores (9, p. 283). It consists simply in finding the difference between each obtained score and the average of the

distribution, and dividing this remainder by the standard deviation of the distribution. Having done this, we were then ready for the following step. We cannot combine reduced scores into a composite score, for this would not yield V and N scores, since we know that the test scores contain both general or central factors and specific factors. We need to separate out the central factor from each of the reduced scores. The resulting scores may then be combined into a composite which will be the score required.

With the correlation data available in Table V, two regression equations were written, one for V, the other for N. This necessitated the calculation of third order partial correlations, the partial standard deviations and regression coefficients. It will be noted that it is the reduced scores from which V and N are to be separated out. Since these reduced scores were obtained by dividing by the obtained standard deviations, they are all of equal weight, and it is therefore unnecessary to use the obtained standard deviations, in computing the partial standard deviations. All standard deviations, then, have a weight of 1.00, and the values obtained from the regression equations can then be applied directly to the reduced scores without overweighting or underweighting. Following this method, then, the regression equations are:

$$V = .5786X_2 + .2912X_3 + .1341X_4 + .0723X_5$$

in which X_2 is Vocabulary, X_3 is Analogies, X_4 is Sentence Completion, X_5 is Disarranged Sentences.

$$N = .3672X_2 + .2858X_3 + .2112X_4 + .1759X_5$$

in which X_2 is Arithmetic Reasoning, X_3 is Number Series Completion, X_4 is Equation Relations, X_5 is Multiplication.

The verbal tests contribute to V in the proportions given above. These proportions are, roughly, in the order given, 8 to 4 to 2 to 1. For the numerical tests, the relative contributions to N are, in the order given, roughly 2.0 to 1.6 to 1.2 to 1.0. In this scheme, the X_5 variable is weighted 1.

If the reduced scores referred to above are multiplied by the weights given in these regression equations, and the resultant scores combined, we shall then have for each subject a V score and an N score, in terms of equated units. This done, we need simply to compute the correlation between the V and N scores, the result being the correlation between V and

N, which is what we set out to find. The necessary computation having been made, it is found that the correlation between the central verbal factor (V) and the central numerical factor (N), is .2625, P. E. .0254. Since it has been shown that the Vocabulary Test is a highly efficient measure of V, it must follow that its correlation with N should approximate the correlation of V with N. This correlation was calculated, and it was found to be .2685. Comparison of this coefficient with that for V and N shows that the difference is very slight.

It will surely be agreed that there is little correspondence between the central verbal factor and the central numerical factor. Yet it has been demonstrated that the verbal tests, taken as a team, satisfy the tetrad difference criterion. This is true, also, for the numerical tests. The lack of correspondence between V & N cannot be due to specific factors, for these have been eliminated. It must be concluded, then, that whatever the nature of V and N, *they are assuredly not the same.* Indeed, it is not unreasonable to ascribe at least part of the correlation of .2625 $\pm$ P. E. 0254 to the fact that in *both* sets the subjects worked with paper and pencil, in the same room, under much the same conditions, *language being used throughout.* This similarity of conditions and partial similarity of content should of itself yield some degree of correlation, i.e., a small general factor, quite independently of similarity of mental function. According to Spearman, the central or general factor is the same throughout all cognitive functions, varying only in the degree to which particular functions draw upon it. It appears, then, that the outcome of this experiment is in conflict with Spearman's interpretation, or else it reduces our 'g' to a relatively small factor, probably verbal, for the most part.

There may be some objection to the foregoing conclusion on the ground that we ought to have calculated the full number of tetrad differences for the entire group of eight tests. It will be recalled that only six tetrad differences were calculated, whereas the full correlation table requires the calculation of 210 tetrad differences. Spearman requires that each and every one of these 210 tetrad differences equal zero before it can be said that the functions tested are divisible into two factors. There is some force in this objection, for it may be that the low correspondence between V and N is due to the presence of group factors, which Spearman interprets as the result of the

overlapping of specific elements. The criterion for the presence of such group factors is the failure of the tetrad difference criterion. This point, then, requires investigation.

E. Group Factors

If the verbal tests contain a group factor which is not present in the numerical tests, the application of Kelley's technique ought to reveal it (15, Prop. 16, p. 69). This proposition reads as follows:

"If the inter-correlations between four variables are such that $t_{1234} = t_{1243}$ and $t_{1342} = 0$, they could conceivably have arisen from four variables x_1, x_2, x_3, x_4 through which was a general factor plus, in addition thereto, a second factor common to x_1 and x_2 or a second factor common to x_3 and x_4." We shall apply this test to a table of intercorrelations involving Vocabulary, Analogies, Arithmetic and Number Series Completion. These are chosen because they are our best tests, in the sense that they have the highest correlations with their central factors. Let us build up the following theoretical scheme, after Kelley: If α = a general factor common to all four variables, β = an additional factor not shared by all of the variables, e = a factor or factors specific to each variable, and C and K = constants which do not change from one variable to another, we may then write:

1. Vocabulary	$= C_1 \alpha_1 + K_1 \beta_1 + e_1$
2. Analogies	$= C_2 \alpha_2 + K_2 \beta_2 + e_2$
3. Arithmetic Reasoning	$= C_3 \alpha_3 + e_3$
4. No. Series Completion	$= C_4 \alpha_4 + e_4$

This gives us a situation in which we have a general factor present in all of the variables, and an additional factor present in only two of the tests, Vocabulary and Analogies. Correlations are obtained by multiplying the variables by one another. The e's are, by definition, specific and therefore uncorrelated, and may therefore be disregarded. The c's and K's are constants, and therefore do not influence the correlations. By multiplication, then, we get the following correlations:*

* Since β does not correlate with anything in the number tests this term disappears from all the correlations except the first.

r_{12} : Vocabulary and Analogies $= \alpha_1 \alpha_2 + \beta_1 \beta_2$
r_{13} : Vocabulary and Arithmetic Reasoning $= \alpha_1 \alpha_3$
r_{14} : Vocabulary and No. Series Completion $= \alpha_1 \alpha_4$
r_{23} : Analogies and Arithmetic Reasoning $= \alpha_2 \alpha_3$
r_{24} : Analogies and No. Series Completion $= \alpha_2 \alpha_4$
r_{34} : Arithmetic and No. Series Completion $= \alpha_3 \alpha_4$

Applying the tetrad difference method, we obtain the following:

$$t_{1234} = \alpha_3 \alpha_4 \beta_1 \beta_2$$
$$t_{1243} = \alpha_3 \alpha_4 \beta_1 \beta_2$$
$$t_{1342} = 0$$

That is to say, the first two tetrads equal each other, while the third equals zero. This scheme has been applied to all the possible combinations of the four variables. In each of these combinations we have assumed the existence of the additional factor in two of the variables, but not in the other two. Finally we assumed an additional factor in the verbal tests, and still another additional factor present in the numerical tests, but absent from the verbal. These combinations have been worked out, in the manner illustrated above, the tetrads which ought to result being presented in Table IX.

TABLE IX

THEORETICAL TETRAD DIFFERENCES WHICH SHOULD RESULT WHEN CERTAIN CONDITIONS CONCERNING GROUP FACTORS ARE SATISFIED

Assuming a Group Factor in:	*Tetrads Should be:*
1. Vocabulary and Analogies	$t_{1234} = \alpha_3\alpha_4\beta_1\beta_2$ $t_{1243} = \alpha_3\alpha_4\beta_1\beta_2$ $t_{1342} = 0$
2. Arithmetic and No. Series Completion	$t_{1234} = \alpha_1\alpha_2\beta_3\beta_4$ $t_{1243} = \alpha_1\alpha_2\beta_3\beta_4$ $t_{1342} = 0$
3. Analogies and Arithmetic	$t_{1234} = 0$ $t_{1243} = \alpha_1\alpha_4\beta_2\beta_3$ $t_{1342} = \alpha_1\alpha_4\beta_2\beta_3$
4. Vocabulary and Arithmetic	$t_{1234} = \alpha_2\alpha_4\beta_1\beta_3$ $t_{1243} = 0$ $t_{1342} = \alpha_2\alpha_4\beta_1\beta_3$
5. Vocabulary and Completion	$t_{1234} = 0$ $t_{1243} = \alpha_2\alpha_3\beta_1\beta_4$ $t_{1312} = \alpha_2\alpha_3\beta_1\beta_4$

6. Analogies and No. Series Completion	$t_{1234} = \alpha_1\alpha_3\beta_2\beta_4$ $t_{1243} = 0$ $t_{1342} = \alpha_1\alpha_3\beta_2\beta_4$
7. Assuming a group factor (β) in Vocabulary and Analogies and an additional group factor (γ) in Arithmetic and No. Series Completion	$t_{1234} = \alpha_3\alpha_4\beta_1\beta_2 + \alpha_1\alpha_2\beta_3\beta_4 + \beta_1\beta_2\beta_3\beta_4$ $t_{1243} = \alpha_3\alpha_4\beta_1\beta_2 + \alpha_1\alpha_2\beta_3\beta_4 + \beta_1\beta_2\beta_3\beta_4$ $t_{1342} = 0$

Having worked out the tetrad differences which are required by theory in the seven situations presented in the foregoing table, it is in order to calculate the actual tetrad differences derivable from the four variables under consideration. These tetrad differences are given below:

$$t_{1234} = .3120$$
$$t_{1243} = .3132$$
$$t_{1342} = .0012$$

It is clear that the first two tetrad differences may be taken as equal to each other, and the third as equal to zero. Within the limits of the probable errors, the situation conforms to Kelley's sixteenth proposition. If we think of the four variables as containing a general factor running through all, and an additional factor present in the two verbal tests, but not present in the numerical tests, we find that the obtained tetrad differences agree with the theoretical prediction. If we assume that the additional factor is present in the numerical tests, but not in the verbal, the facts also fit the theory. Lastly, if we assume an additional factor present in each of the two types of tests, there is still agreement with the theory. Of particular significance is the fact that no group factor cuts across the two forms of tests. There is no single group factor present in both the verbal and the numerical tests. This does not mean that we can be certain of two group factors, though two may exist. It is practically certain, however, that at least *one* group factor exists, either in the verbal or in the numerical tests. Complete certainty is not yet possible, since the converse of Kelley's proposition has not been proved. But the probabilities are over-whelmingly in favor of the presence of at least one group factor, confined to either the verbal or the numerical tests.

These conclusions are in agreement with the low correlation between the verbal and numerical abilities, as expressed in the

coefficient of .2625. Our original conclusions appear to be substantiated. Whatever the nature of V and N, the two are not the same. If the group factor is in the verbal tests, it cannot be due to the overlapping of specific factors, for the entire battery of verbal tests satisfies the tetrad difference criterion, i.e., all of the tetrad differences are zero, within the limits of probable error. The same observation applies equally if the group factor be located in the numerical tests. It still holds if there be two group factors, one in each battery of tests. The tetrad differences are not zero when the verbal and numerical tests are taken together, and this, according to Spearman, proves the existence of group factors. Since each battery, taken alone, satisfies the criterion, and since the criterion fails only when we cut across the boundary between the two batteries, it follows that the group factor should also cut across the boundary. But it has just been shown that this is not the case. In Table IX we have shown the statistical results which should be found if the group factor cuts across the two types of tests. These results, however, are not found. 'The observed facts will agree only with the theoretical predictions which are based on the assumption of a verbal factor, or a numerical factor, or both. There is no agreement with the assumption of a verbal-numerical factor, over and above 'g'.

3. *Relation of V and N to Other Abilities*

It has been shown that V and N are distinctive in the sense that their correspondence is quite low. There arises the further question as to whether they are also distinguishable from other abilities. It is possible that they may be nothing more than measures of what may be called 'general ability', and not distinctive traits. The answer to this question may be found by testing our group in other abilities, and then correlating these abilities with V and N. One such test has been made, in the field of memory.

Several weeks prior to the major experiment, a paired associates test in memory was given to the same group. The test was given in laboratory, in groups of 18 subjects. Absences reduced the size of the group from 210 to 187. The material was paired as follows:

pictures and pictures
words and words
words and nonsense syllables
geometrical forms and digits

Two forms of the test were given visually, twenty pairs being contained in each of the forms. In each form, there were five pairs of each type as described above, making a total of ten pairs of each kind, and a total of forty pairs in all. The pictures were taken from periodicals. They were all in black and white, all of the same size, and all presented on cards measuring nine inches by four inches. They were pictures of common objects, such as a piano, a pencil, a watch, a dog, a boy, a bed, a shoe, and the like. The words were common four-letter words, without obvious associations. A few examples are: foot, past; four, bade; foal, high; trim, lark; fish, show. Some examples of the words and nonsense syllables are given: ring, bef; rake, dro; duck, orp; safe, ilt; the rest were of the same form. The geometrical patterns and digits were paired after the following manner; triangle, 56; oval, 27; diamond, 12, etc. All pairs were affixed to cards of the size mentioned for the pictures. The words, nonsense syllables and digits were all of the same size, printed on white cards, in black ink. The geometrical forms were approximately equivalent in area, inked in black and affixed to white cards. The exhibits were presented in irregular order, each being shown during four beats of a metronome set at 60.

The test was for immediate retention. The cards were presented a second time, the right side being in each case covered, only the left side being visible to the subjects. The subjects were required to state, in writing, what had been shown on that side of the card which was now covered. Both forms of the test were given during the same class period. Scoring was in terms of the number of items remembered. The results yielded the following table of correlations:*

TABLE X

CORRELATIONS BETWEEN VERBAL, NUMBER AND MEMORY TESTS

	Memory
Memory	.750 (self-correlation)
V	.114
N	.171
Synonyms	.070
Analogies	.115
Sentence Completion	.197
Disarranged Sentences	.067
Arithmetic Reasoning	.133
Number Series Completion	.164
Equation Relations	.086
Age	.077

* The writer is indebted to Mr. Murray Poliwinchik for the calculation of these coefficients.

The correlation between memory and memory (.750) is the reliability coefficient, found by correlating the two forms against each other, and applying the Spearman-Brown Formula. The other correlations were found by using, for each subject, his total score for each form of the memory test, and then combining these into a single score for the entire memory test, whose reliability has been given above. The reliability coefficient is not as high as might be desired, but it is, nevertheless, of reasonable magnitude. A practice effect is indicated by increased scores for the second form of the test. It is felt that the reliability of the test might be higher if preliminary practice had been given.

In no case is the correlation between memory and the other variables as high as .20. This points very clearly to the relative independence of the verbal and numerical abilities from memory, as measured by the test employed. This is the more striking, in view of the fact that the tests for retention involved words and numbers. Even the pictures and geometrical forms were probably dealt with verbally during the course of the experiment. In any case, the responses in the test for retention were written, and always in the form of a word, nonsense syllable or number. Even this community of form of content did not suffice to give appreciable correlations. One may, perhaps, be pardoned a guess that whatever correlation is present is due, in part at least, to similarity of content. It is conceivable that a test for retention which excluded words, syllables and digits might have yielded even a lower correlation with V and N.

4. *Relations of Verbal and Numerical Abilities to Scholastic Records.*

The writer believes that, so far as this study is concerned, the relative independence of V and N has been established. It was thought desirable, nevertheless, to compare the results with the scholastic records of the subjects. It is reasonable to expect some correspondence between performance in the tests and the work done by the students in their college courses. The validity of the conclusions already offered does not depend upon the existence of such correspondence, for it is well known that scholastic grades in college are by no means a consistently accurate criterion of the students' abilities. Many factors other than scholastic ability play a part in the determination

of grades. On the other hand, it is felt that the tests used in this study have yielded a fair gauge of ability in the functions tested, principally because of the novelty of the situation and the fact that the work was difficult enough to be challenging to the subjects. A close correspondence between performance in the tests and scholastic records is not to be expected. There should, nevertheless, be a trend, however slight, in the direction of correspondence. At any rate, it was thought that the comparison might be interesting, and it was, therefore, made.*

Scholastic records for the group used in the retention test were gathered from the files of the Registrar of the College. The record of one man was missing, so that the group was reduced in number from 187 to 186. The comparison was made in three ways:

1. The subjects' grades in literary and in science courses were collected. Literary courses were defined as those which involved direct study of language and literature. All courses in English, French, German, Spanish, Italian, Latin and Greek were included in this group. For the science group, grades were taken for all courses in Mathematics, Physics, Chemistry, Astronomy, Engineering and Biology. Letter grades were converted into numerical values by the following arbitrary weighting:

Grade A = 4 points
Grade B = 3 points
Grade C = 2 points
Grade D = 1 point.

The number of credits received by the subject in a given course was multiplied by the value given above. This yields a weighted value in points for each course. These weighted values were summed separately for the literary courses and for the science courses, giving, for each man, a total score for literature and for science. Each man's literature score was then divided by his total number of semester credits in literary courses, the result being an average value for the literary courses. The same procedure was followed with respect to the science courses. The subjects were then divided into three groups, based upon their average grade values. Group A stands high-

* The writer acknowledges his indebtedness to Mr. Murray Poliwinchik, who gathered the scholastic records from the files of the Registrar of the College. Mr. Poliwinchik also made most of the calculations recorded in the tables following.

est, Group B is the middle group, and Group C is the lower group. There are, of course, three groups under the heading of literature, and three under the heading of science. For each of these six groups, the average V and N scores were computed. It will be recalled that V and N scores were expressed in terms of sigma values. This means that they are fractional values, some positive and some negative. For the sake of simplicity of presentation, these sigma scores were converted in the manner described by Hull (9). Assuming a normal distribution with a mean of 50, a range of 0 to 100, and a standard deviation of 14, sigma scores may be converted readily into values in such a distribution. This, accordingly, was done, and the verbal and numerical scores given below are expressed in terms of such values.

TABLE XI-a

Comparison of V and N Scores With Ability in Literary and Science Courses

* Group A = highest grades. Group B = middle grades. Group C = lowest grades.

LITERARY COURSES				*SCIENCE COURSES*			
Group	*N***	*Verbal*	*Numerical*	*Group*	*N***	*Verbal*	*Numerical*
A	57	56.5	51.3	A	53	53.1	54.1
B	90	49.2	48.3	B	85	50.2	50.4
C	39	43.8	50.6	C	48	49.6	46.7

Analyzing the data for the literary courses, it is found that Group A, which is of superior ability in literary courses, has a higher average V score than does either of the other groups. The V score for Group A is also higher than the N score for the same group. The middle group, B, has a higher V score than does Group C, but its V score is only slightly higher than is its N score. This is not surprising, for it is not to be expected that a group which is mediocre in literary courses should be better in these than in science courses. Group C has the lowest V score, which was to be expected. Since this group is inferior in literary courses, one need not be surprised to find its N score higher than its V score; and this is precisely what is found. These results are clearly indicative of a trend towards correspondence between V and ability in literary courses.

* A, B, C are not grades. These letters merely designate the high, low and middle groups.

** N here refers to population.

Turning to the science group, a similar trend is noted. Group A has a somewhat higher score for N than for V. It is also the best group in N. Group B has a higher N score than does Group C, but, as before, there is practically no difference between V and N for this group. Group C has the lowest N score, and its V is higher than its N. The trend of results for the science courses parallels that for the literary courses quite closely. Comparing the literary with the science group, the expected results are found. Group A (literary) has a higher V score than does Group A (science), but a lower N score. Group C (literary) being inferior in literary courses, has a lower V score than does Group C (science), but a higher N score. Since the B groups are mediocre, there is no reason to expect any advantage in either direction. The B (science) group does have a higher V score than does the B (literary) group, but it has, also, a higher N score. Though these differences are small, they nevertheless constitute a departure from the otherwise uniform trend in the expected direction. This deflection, however, is not adverse to the general drift. It is simply outside the current. On the whole, the comparison points to some agreement between performance in the tests and performance in college work. This agreement, however, is not close.

2. Comparison with scholastic ability was made in another way. The group of 186 men was divided into two classes, the literary and the scientific. The literary group includes those men who have a greater number of semester credits in literary courses than in science courses. The scientific group includes those who have a preponderance of credits in science courses. Each of these groups was sub-divided into three parts, A, B and C, based upon the number of credits earned. It was thought that mere exposure to literary courses or to science courses ought to be reflected in the V and N scores. The C group includes those who had earned from 11 to 24 credits; the B group, from 25 to 38 credits; the A group, from 39 to 52 credits. The division was made in this manner for both the literary and the scientific men.

The average V and N scores were computed for each of the six sub-groups, with results as shown in Table XI-b.

The trend is fairly marked. The A and B literary groups have noticeably higher scores for V than for N. The A and B science men have distinctly higher scores for N than for V.

TABLE XI-b
COMPARISON OF V AND N SCORES WITH TRAINING IN LITERARY AND SCIENCE COURSES
A, B, C, are defined as in Table XI-a.

LITERARY COURSES				*SCIENCE COURSES*			
Group	*N*	*Verbal*	*Numerical*	*Group*	*N*	*Verbal*	*Numerical*
A	26	54.4	48.4	A	33	48.0	52.8
B	40	54.5	48.6	B	39	48.2	52.2
C	30	46.2	46.5	C	18	52.3	48.8

The C literary men are very slightly superior in N than in V, while the C science men have higher V scores than N scores. These results agree very well with the hypothesis that mere exposure to literary courses ought to be reflected in higher V scores than N scores, regardless of grades. Similarly, a preponderance of work in science courses results in better N scores than V scores regardless of grades. The outcome is the same when the literary men are compared with the science men. The V scores for the former are higher than those for the latter, at least for the A and B groups. Conversely, the N score for the science men is higher than that for the literary men, Group C excepted. Though the C science men had more work in science than in literature they nevertheless appear to have had too little of science to develop superiority in N. It seems probable that training in other subjects, most of them involving verbal elements, had more effect than their relatively meagre training in scientific work.

3. Comparison of V and N with scholastic records was made in still a third way. The entire group was divided into two classes as before, literary and scientific men, based upon the field in which a preponderance of work was done. Instead of dividing each of these two classes into three sub-groups, based upon the number of credits earned, the division was made on the basis of average grades. Thus, the literary men were divided into three groups, A, B and C, the A group including the literary men who had the highest grade in literature, the C group including those literary men who had the lowest grades, the B group being intermediate. A similar division was made for the science men. It was thought that, of those men who had done a preponderance of work in literature, those who had earned the best grades ought to have the best V scores, while the best science men ought to have the best N scores. The results are shown in Table XI-c.

It is found that for the literary men, the V scores progress

TABLE XI-c

COMPARISON OF V AND N SCORES FOR LITERARY AND SCIENCE MEN WITH THEIR GRADES IN THEIR RESPECTIVE FIELDS

LITERARY MEN				*SCIENCE MEN*			
Group	*N*	*Verbal*	*Numerical*	*Group*	*N*	*Verbal*	*Numerical*
A	30	53.2	52.6	A	33	50.2	53.6
B	41	47.3	47.3	B	43	48.5	51.7
C	25	44.4	47.8	C	14	46.7	46.9

downward as we proceed from Group A to Group C. A similar result is found for the science men. Their N scores become lower as we proceed from Group A to Group C. It is curious, however, that in the literary group, differences between the V and N scores are practically non-existent, excepting for the C men, and here the difference is, as before, in the reverse direction. On the whole, it appears justifiable to conclude that V and N scores are, in slight degree, in agreement with the quality of work done in corresponding courses in college. It would be premature to say that there is a close interdependence. It will be noticed that division of the experimental group into sub-groups resulted in very small populations, from which it would be rash to draw any far-reaching conclusions. There is, undoubtedly, a drift in the direction of agreement, but the outcome is, at best, only suggestive.

It would, indeed, be somewhat surprising if the agreement were much closer than has been found in this study. After all, there are many courses, other than the strictly literary, which are primarily, if not wholly, verbal in content. On the other hand, all courses in science, even mathematics, involve considerable verbal elements.* Where there is so much overlapping, it would be astonishing if clear-cut distinctions could be made. It would seem worth-while to submit the question to carefully controlled experiment. It is believed that the evidence here adduced is sufficient to warrant the tentative hypothesis that V and N are, in part, determined by training in verbal and numerical tasks. The most convincing evidence is contained in Table XI-b, from which, upon comparison with the other tables, it appears that there is closer correspondence of V and N with amount of work done, rather than quality of work. This last fact is not astonishing, for the unreliability of college grades is notorious.

* It is suggested that science men require reading comprehension, and this would tend to obscure the difference between verbal ability and numerical ability. On the other hand, numerical ability is not requisite for success in V.

V

SUMMARY

1. The problem was concerned with the inter-relationship existing between verbal and numerical abilities; the relation of these to other abilities; and the determination of the best instruments for measuring verbal and numerical abilities.

2. The subjects were 210 men representing a minimum of diversity with respect to education, ancestry, economic and social status, and age.

3. Nine tests were given, five verbal and four numerical, group method used.

4. General factors for verbal ability (V) and for numerical ability (N) were found.

5. Verbal and numerical ability were found to have little in common, the correlation between V and N being .2625.

6. It was established that either V or N contains something which is not present in the other.

7. The paired associates test for immediate retention was given and it was found that V and N, besides being differentiated from each other, are also distinguishable from retention. V and N are not, therefore, merely measures of general ability.

8. There is a noticeable but not marked correspondence between verbal ability and performance in literary courses in college. The same condition holds with respect to numerical ability and the performance in science courses, including mathematics.

9. The best single test of verbal ability is a vocabulary test. The best single test of numerical ability is the Arithmetic Test. For each ability the entire battery of tests is more efficient than any single test.

VI. BIBLIOGRAPHY

1. Anderson, R. C. 1925. A critical analysis of test-scoring methods. Archives of Psychology No. 80.
2. Burt, C. 1916. Distribution and relation of educational abilities. London County Council.
3. Burt, C. 1922. Mental and scholastic tests. London. King.
4. Bonser, F. G. 1910. The reasoning ability of children of the 4th, 5th, and 6th grades. T. C. Cont. to Education No. 37.
5. Bronner, A. F. 1917. The psychology of special abilities and disabilities. Boston, Little, Brown & Co.
6. Davey, C. M. 1926. Comparison of group verbal and pictorial tests of intelligence. Brit. J. Psych. 17, 27-48.
7. Dodd, S. C. 1928. The theory of factors. II. Psych. Rev. 35, No. 4, 268-279.
8. Dodd, S. C. 1928. The theory of factors. I. Psych. Rev. 35, No. 3, 211-234.
9. Garrett, H. E. 1926. Statistics in psychology and education. New York. Longmans, Green & Co.
10. Gates, A. I. 1922. Correlations of achievement in school subjects with intelligence tests and other variables. J. Ed. Psych. 13, 129-139.
11. Gates, A. I. 1922. Psychology of reading and spelling. T. C. Cont. to Educ. No. 129.
12. Gates, A. I. 1928. The improvement of reading. New York. Macmillan.
13. Hollingworth, L. S. and Winford. 1918. The psychology of special disability in spelling. T. C. Cont. to Educ. No. 88.
14. Holzinger, K. J. 1929. On tetrad differences with overlapping variables. J. Ed. Psych. 20, No. 2, 91-97.
15. Kelley, T. L. 1928. Crossroads in the mind of man. Stanford University, Stanford University Press.
16. Kelley, T. L. 1924. Statistical method. New York. Macmillan.
17. Spearman, C. 1927. The abilities of man. New York. Macmillan.
18. Spearman, C. 1912-1913. Correlation of sums and differences. Brit. J. Psych. 5, 417-426.
19. Symonds, P. M. 1923. Special abilities in algebra. Teachers' College, Columbia University, New York.
20. Terman, L. M. 1916. The measurement of intelligence. New York. Houghton, Mifflin Co.
21. Thomson, G. H. 1927. The tetrad difference criterion. Brit. J. Psych. 17, 235-255.
22. Thorndike, E. L. 1927. The measurement of intelligence. New York. Teachers' College, Columbia University.
23. Thorndike, E. L. 1921. Intelligence and its measurement: A symposium. J. Ed. Psych. 12, 127-129.
24. Wilson, E. B. 1929. Comment on Professor Spearman's note. J. Ed. Psych. 20, No. 3, 217-223.

BF 21 A7 no.107 40955

Schneck, Matthew M. R.

The measurement of verbal and numerical abilities.

Date	Issued to

PSYCH

BF21
A7
no.107

Schneck, Matthew Maximilian Rupprecht, 1898-
The measurement of verbal and numerical abilities, by Matthew M. R. Schneck. New York, 1929.
49 p. 25 cm. (Archives of psychology ... no. 107)

40955

PSYCH

MBNU 17 AUG 77 3194520 NEDZbp 29-23507

ARCHIVES OF PSYCHOLOGY

List of numbers, continued from inside front cover

63. Race Differences in Inhibition: ALBERT L. CRANE. $1.50.
64. Individual Differences in Incidental Memory: SADIE MYERS SHELLOW. $1.25.
65. Character Traits as Factors in Intelligence Test Performance: WILLIAM M. BROWN. $1.25.
66. A Study of the Sexual Interest of Young Women: F. I. DAVENPORT. $1.25.
67. The Psychology of Confidence: WM. CLARK TROW. $1.25.
68. Experimental Studies of College Teaching: HAROLD E. JONES. $1.25.
69. The Influence of Treatment for Intestinal Toxemia on Mental and Motor Efficiency: ALICE E. PAULSEN. $1.00.
70. A Study of Suggestibility of Children: MARGARET OTIS. $1.50.
71. The Value of Praise and Reproof as Incentives for Children: ELIZABETH B. HURLOCK. $1.00.
72. Attention and Interest in Advertising: HOWARD K. NIXON. $1.25.
73. An Experimental Study of Thinking: EDNA HEIDBREDER. $1.75.
74. Estimation of Time: ROBERT AXEL. $1.00.
76. Measurement of Emotional Reactions: DAVID WECHSLER. $1.75.
77. Tested Mentality as Related to Success in Skilled Trade Training: THEODORA M. ABEL. $1.25.
78. Aggressive Behavior in a Small Social Group: E. M. RIDDLE. $1.75.
79. The Memory Value of Advertisements: EDITH R. BRANDT. $1.25.
80. A Critical Examination of Test-Scoring Methods: ROSE G. ANDERSON. $1.00.
81. Thermal Discrimination and Weber's Law: ELMER A. K. CULLER. $1.75.
82. A Correlational Analysis of Typing Proficiency: LUTON ACKERSON. $1.50.
83. Recall as a Function of Perceived Relations: CORA B. KEY. $1.25.
84. A Study of the Consistency of Rate of Work: CONSTANCE E. DOWD. $1.00.
85. An Experimental Investigation of Recovery from Work: S. L. CRAWLEY. $1.25.
86. Facilitation and Inhibition: THOMAS N. JENKINS. $1.00.
87. Variability of Performance in the Curve of Work: JAMES D. WEINLAND. $1.00.
88. A Mental Hygiene Inventory: S. DANIEL HOUSE. $1.50.
89. Mental Set and Shift: ARTHUR T. JERSILD. $1.25.
90. An Experimental Investigation of Rest Pauses: CHARLES W. MANZER. $1.25.
91. Routine and Varying Practice as Preparation for Adjustment to a New Situation: LELAND W. CRAFTS. $1.00.
92. The Development of a Standardized Animal Maze: L. H. WARNER and O. J. WARDEN. $1.25.
93. An Experimental Study of Speed and Other Factors in "Racial" Differences: OTTO KLINEBERG. $1.50.
94. The Relation of Reaction Time to Measures of Intelligence, Memory, and Learning: VERNON W. LEMMON. 80 cents.
95. Is the Latent Time in the Achilles Tendon Reflex a Criterion of Speed in Mental Reactions? GEORGE H. ROUNDS. $1.25.
96. The Predictive Value of Certain Tests of Emotional Stability as Applied to College Freshmen: EDWIN G. FLEMMING. $1.00.
97. A Vocabulary Information Test: ANGELINA L. WEEKS. $1.00.
98. The Effect of Various Temporal Arrangements of Practice on the Mastery of an Animal Maze of Moderate Complexity: SIDNEY A. COOK. 80 cents.
99. Recognition Time as a Measure of Confidence: GEORGENE HOFFMAN SEWARD. $1.00.
100. Precision and Accuracy: GEORGE W. HARTMANN. 80 cents.
101. A Group Test of Home Environment: EDITH MARIE BURDICK. $1.50.
102. The Effect of Changed Material on Ability to Do Formal Syllogistic Reasoning: MINNA CHEVES WILKINS. $1.25.
103. The Effect of Incentives on Accuracy of Discrimination Measured on the Galton Bar: HUGHBERT C. HAMILTON. $1.25.
104. An Investigation into the Validity of Norms with Special Reference to Urban and Rural Groups: MYRA E. SHIMBERG. $1.25.
105. Blood Pressure Changes in Deception: MATTHEW N. CHAPPELL. 80 cents.

In addition to the numbers of the Archives, the following monographs are to be obtained from us:

The Psychology of Association: FELIX ARNOLD. 50 cents.
The Measurement of Variable Quantities: FRANZ BOAS. 50 cents.
The Diurnal Course of Efficiency: HOWARD D. MARSH. 90 cents.
The Time of Perception as a Measure of Differences in Sensations: VIVIAN ALLEN CHARLES HENMON. 60 cents.
Interests in Relation to Intelligence: LOUISE E. POULL. $1.00. Reprinted from *Ungraded.*
The Conditioned Pupillary and Eyelid Reactions: HULSEY CASON. $1.00. Reprinted from the *Journal of Experimental Psychology.*

Date Due

DIRECTORY OF
AMERICAN PSYCHOLOGICAL PERIODICALS

AMERICAN JOURNAL OF PSYCHOLOGY—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages ann. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY—Worcester, Mass.; Clark University Press. Subscription $7.00. 700 pages ann. Edited by Carl Murchison and an international coöperating board. Quarterly. Child behavior differential and genetic psychology. Founded 1891.

PSYCHOLOGICAL REVIEW—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 480 pages annually.
By-monthly. General. Founded 1894. Edited by Howard C. Warren.

PSYCHOLOGICAL MONOGRAPHS—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per volume. 500 pages. Founded 1895.
Edited by Raymond Dodge.
Published without fixed dates, each issue one or more researches.

PSYCHOLOGICAL INDEX—Princeton, N. J.; Psychological Review Company.
Subscription $4.00. 300-400 pages. Founded 1895. Edited by W. S. Hunter.
An annual bibliography of psychological literature.

PSYCHOLOGICAL BULLETIN—Princeton, N. J.; Psychological Review Company.
Subscription $6.00. 720 pages annually. Psychological Literature.
Monthly. Founded 1904. Edited by Samuel W. Fernberger.

JOURNAL OF PHILOSOPHY—New York; 515 W. 116th St.
Subscription $4.00. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider.

TRAINING SCHOOL BULLETIN—Vineland, N. J.; The Training School.
Subscription $1.00. 160 pages ann. Edited by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

ARCHIVES OF PSYCHOLOGY—Columbia University, N. Y.; Archives of Psychology.
Subscription $6.00. 500 pp. per vol. Founded 1906. Edited by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY—Albany, N. Y.
Sub. $5.00. Boyd Printing Co. Edited by Morton Prince, in cooperation with Floyd H. Allport.
Quarterly. 432 pages annually. Founded 1906. Abnormal and social.

PSYCHOLOGICAL CLINIC—Philadelphia; Psychological Clinic Press.
Subscription $2.50. 288 pages. Edited by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology hygiene.

COMPARATIVE PSYCHOLOGY MONOGRAPHS—Baltimore; Williams and Wilkins Co.
Subscription $5.00. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

PSYCHOANALYTIC REVIEW—Washington, D. C.; 3617 10th St., N. W.
Subscription $6.00. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Edited by W. A. White and S. E. Jelliffe.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY—Princeton, N. J.
Psycological Review Company. 500 pages annually. Experimental.
Subscription $6.00. Founded 1916. Bi-monthly. Edited by Madison Bentley.

JOURNAL OF APPLIED PSYCHOLOGY—Bloomington, Ind.; Indiana University Press.
Subscription $4.00. 400 pages annually. Founded 1917.
Quarterly. Edited by James P. Porter and William F. Book.

JOURNAL OF COMPARATIVE PSYCHOLOGY—Baltimore; Williams and Wilkins Company.
Subscription $5.00. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

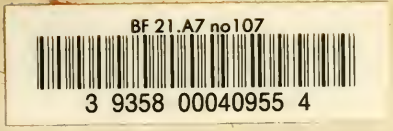
GENETIC PSYCHOLOGY MONOGRAPHS—Worcester, Mass.; Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison and an international cooperating board. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

PSYCHOLOGICAL ABSTRACTS—Princeton, N. J.; American Psycholigical Association.
Subscription $6.00. 600 pages annually. Edited by W. S. Hunter.
Monthly. Abstracts of Psychological literature. Founded 1927.

THE PERSONNEL JOURNAL—Baltimore; William and Wilkins Co.
Subscription $5.00. 500 pp. Founded 1922.
Bi-monthly. Edited by Walter V. Binghan.

JOURNAL OF GENERAL PSYCHOLOGY—Worcester, Mass.; Clark University Press. Subscription $7.00. 600-700 pages annually. Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

ARCHIVES OF PSYCHOANALSIS—New York; 2 East 65th St. Subscription $20.00. 1000 pages ann. Ed. by L. Pierce Clark. Quarterly. Detailed analysis of cases of narcistic neuroses and psychoses showing psychoanalytic technic. Foreign abstracts and translations.



This book may be kept.....one.....weeks.

A fine of two cents will be charged for each day books or magazines are kept overtime.

Two books may be borrowed from the Library at one time.

Any book injured or lost shall be paid for by the person to whom it is charged.

No member shall transfer his right to use the Library to any other person.

BF 21.A7 no107
3 9358 00040955 4